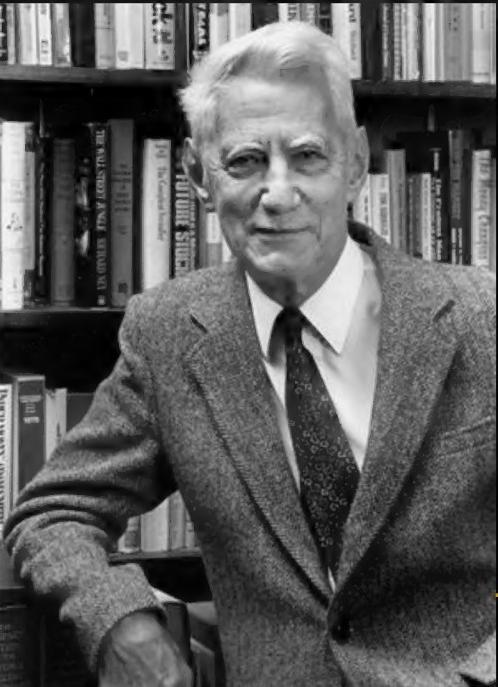


# 信息论基础

## —信息论与编码



# 伟大的科学家—香农

- “通信的基本问题就是在一点重新准确地或近似地再现另一点所选择的消息”。
- 这是数学家香农（**Claude E.Shannon**）在他的惊世之著《通信的数学理论》中的一句名言。正是沿着这一思路他应用数理统计的方法来研究通信系统，从而创立了影响深远的信息论。

# 伟大的科学家—香农

- 香农于1916年4月30日出生于美国密歇根州的Petoskey，并且是爱迪生的远亲戚。1936年毕业于密歇根大学并获得数学和电子工程学士学位，在那里他遇到了离散数学奠基人乔治·布尔，上过他的课程。1940年获得麻省理工学院（MIT）数学博士学位和电子工程硕士学位。1941年他加入贝尔实验室数学部，工作到1972年。1956年他成为麻省理工学院（MIT）客座教授，并于1958年成为终生教授，1978年成为名誉教授。香农博士于2001年2月24日去世，享年84岁。
- 香农于1940年在普林斯顿高级研究所（The Institute for Advanced Study at Princeton）期间开始思考信息论与有效通信系统的问题。经过8年的努力，香农在1948年6月和10月在《贝尔系统技术杂志》（Bell System Technical Journal）上连载发表了他影响深远的论文《通讯的数学原理》。1949年，香农又在该杂志上发表了另一著名论文《噪声下的通信》。在这两篇论文中，香农阐明了通信的基本问题，给出了通信系统的模型，提出了信息量的数学表达式，并解决了信道容量、信源统计特性、信源编码、信道编码等一系列基本技术问题。两篇论文成为了信息论的奠基性著作。

# 香农的贡献

- 1948年发表“通信的数学理论”，标志着信息论的诞生；
- 1949年发表“噪声下的通信”，为信道编码奠定理论基础
- 1949年发表“保密通信的信息理论”，首先用信息论的观点对信息保密问题作了全面的论述；
- 1959年发表“保真度准则下的离散信源编码定理”-提出信息率失真理论，为信源压缩编码研究奠定理论基础；
- 1961年发表“双路通信信道”，开拓了多用户信息理论（网络信息论）的研究；

# 信息论基础的重要性

- 信息论是信息科学和技术的基本理论,信息科学大厦的地基;
- 没有信息论的基础,从事通信与信息领域的研究和创新是不可能的事情;
- 总之,信息论是高层次信息技术人才必不可少的基础知识。

课程类型：专业选修课

学时：54

教材：《信息论与编码》，陈运，周亮，陈新，电子工业出版社，2002

参考教材：

① <<信息论基础>> (Elements of Information Theory, 英文影印), Thomas M.Cover, Joy A.Thomas ,清华大学出版社, 2003

② <<信息论与编码理论>>(The Theory of Information and Coding,英文影印版)(第2版), Robert J.McEliece, 电子工业出版社, 2005.3

考核：平时成绩40%（作业、考勤、期中测验）

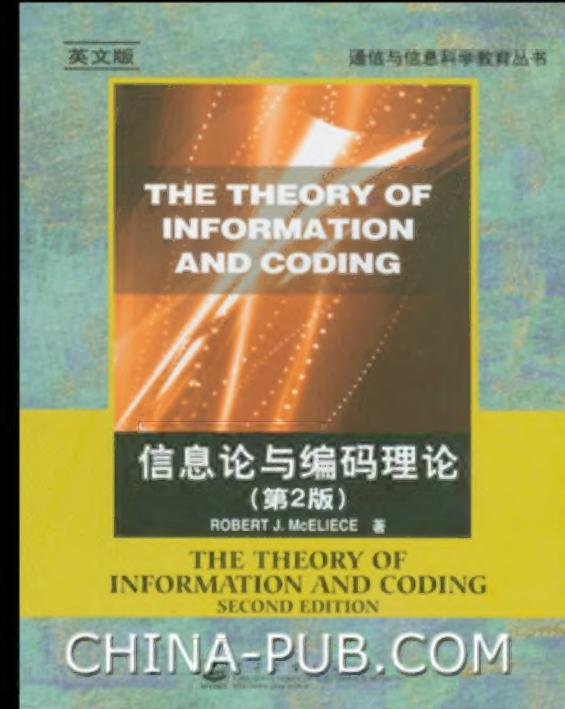
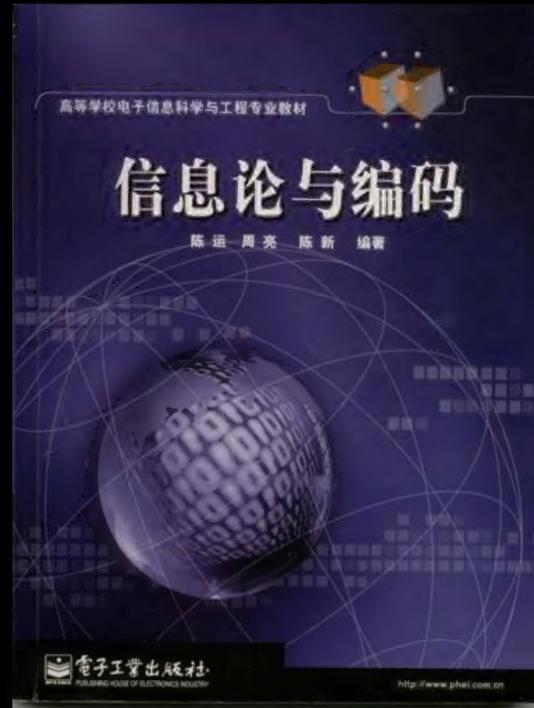
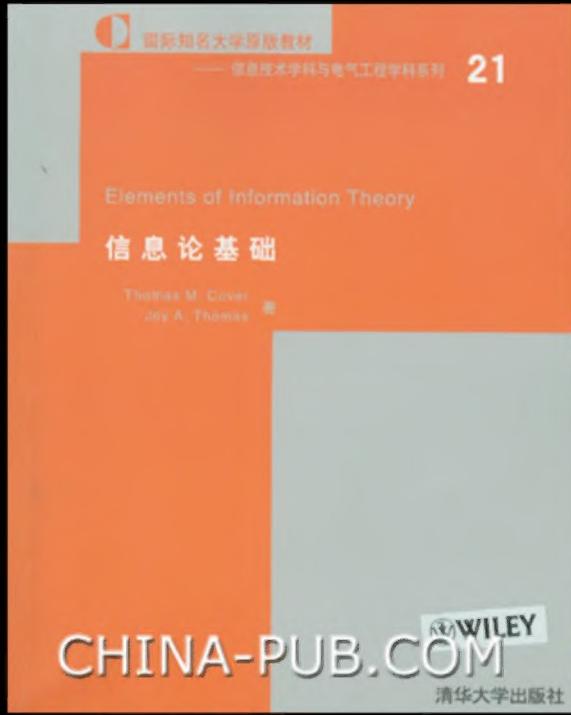
期末考试60%（闭卷）

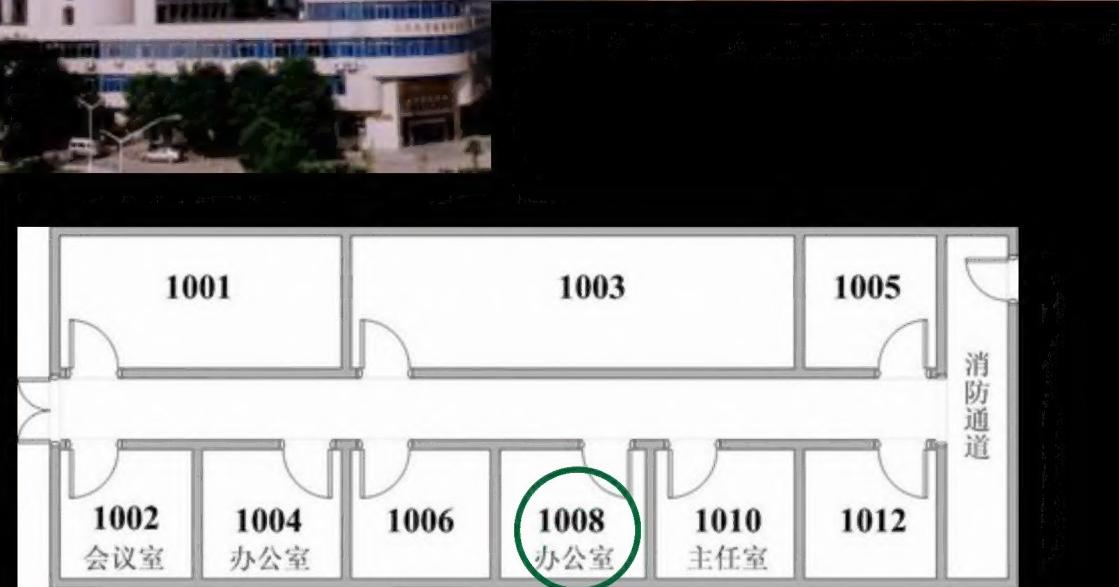
## 上课地点、时间:

- 星期一, 第3-4节 (9: 50-11: 25 AM), 测绘校区, 附1-603 (Every week 1-18)
- 星期三, 第3-4节 (9: 50-11: 25 AM), 测绘校区, 附1-603 (Odd week 1-18)

## 答疑地点、时间:

- 地点: 教学实验大楼十楼信号处理实验室 - 1008;
- 时间: 星期四, 2: 30-5: 00, PM
- 其它答疑方式:
  - 电子邮件: [yw@eis.whu.edu.cn](mailto:yw@eis.whu.edu.cn);
  - 办公电话: 027-68778181





Here you  
can find me!

E-mail: [yw@eis.whu.edu.cn](mailto:yw@eis.whu.edu.cn)  
Tel(O): 027-68778181

# 课程基本內容

## 1 概论（2学时）

1.1 信息的一般概念

1.2 信息的分类

1.3 信息论的起源、发展及研究内容

## 2 信源熵（12学时）

2.1 单符号离散信源

2.2 多符号离散平稳信源

2.3 连续信源

2.4 离散无失真信源编码定理

## 3 信道容量（8学时）

3.1 信道的数学模型和分类

3.2 单符号离散信道的信道容量

3.3 多符号离散信道

3.4 多用户信道

3.5 连续信道

3.6 信道编码定理

## 4 信息率失真函数 (6学时)

- 4.1 基本概念
- 4.2 离散信源的信息率失真函数
- 4.3 连续信源的信息率失真函数
- 4.4 保真度准则下的信源编码定理

## 5 信源编码 (6学时)

- 5.1 香农编码
- 5.2 费诺编码
- 5.3 哈夫曼编码
- 5.4 游程编码
- 5.5 冗余位编码

## 6 信道编码 (16)

- 6.1 信道编码的概念
- 6.2 线性分组码
- 6.3 循环码
- 6.4 卷积码

## 7 保密信息理论 (4)

- 7.1 密码学的基本知识
- 7.2 数据加密标准DES
- 7.3 国际数据加密算法
- 7.4 公开密钥加密原理

# 教学目标

- 充分理解、熟练掌握教材的内容
  - 熟练掌握基本的信息理论概念和定理
  - 熟练掌握通信与信息工程中基本研究对象的数学描述
- 通过学习和习题练习，具备一定的解决问题分析问题的能力
- 掌握一定的科学思想方法(用信息论的思想和观点)

# 学习方法

本课程以概率论为基础，数学推导较多，学习时主要把注意力集中到概念的理解上，不过分追求数学细节的推导。学习时一定要从始至终注意基本概念的理解，不断加深概念的把握。学习时注意理解各个概念的“用处”，结合其他课程理解它的意义，而不要把它当作数学课来学习，提倡独立思考，注重思考在学习中的重要性。

# 对学习者的要求

- 三个重要环节
  - 课前预习
  - 课上认真听讲
  - 课后认真复习消化、作业
- 经常进行阶段复习
  - 掌握知识的窍诀：反复思维实践
- 欢迎大家及时提出反馈意见，以便更好的进行互动教学。

# 其他约定

- 不得迟到、早退、缺课，有事请假
- 上课时请关闭手机
- 作业不得用纸片信纸之类，必须使用作业本
- 迟交的作业及纸片做的作业恕不修改，只作记录

# 计分方式

- 最后期终考试成绩占60%
- 平时成绩占20%（作业20%+考勤10%）
  - 作业：一次不交扣1分，无故缺席一次扣1分，迟到一次扣0.5分
  - 手机声响扣1分
  - 严重违反课堂纪律，视情节轻重扣分
- 期中考核占10%

# 1 概论

## 1.1 信息的一般概念

## 1.2 信息的分类

## 1.3 信息论的起源、发展及研究内容

# 1.1 信息的一般概念

- (1) 当代文明的三大支柱
- (2) 信息是什么？
- (3) 香农和维纳对信息的定义
- (4) 信息的定义和性质

# （1）当代文明的三大支柱

信息科学、材料科学和能源科学一起被称为当代文明的“三大支柱”。是科学历史上三个最重要的基本概念。

- 没有物质的世界是虚无的世界；
- 没有能源的世界是死寂的世界；
- 没有信息的世界是混乱的世界。



## 「(2) 信息是什么？」

- ①从通信谈起—通信系统的基本概念
- ②信息科学的有关概念

- 信息与消息和信号的区别
- 各种信息的概念
- 信息的直观认识

# ①通信系统的基本概念

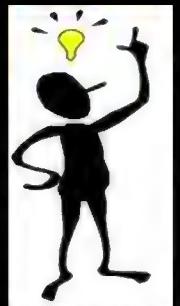
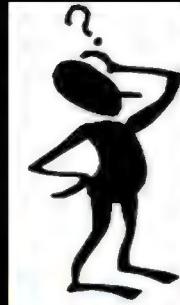
## ■ 从通信谈起：

- 通信的目的是传输信息，即传递消息。
- 近代的通信特指“电通信”。
- 通信系统的主要任务是借助电信号这种载体来传输消息。
- 信息（Information）、消息（Message）、信号（Signal）就成为人们最常使用的3个基本术语



- 它们的确切定义及基本含义是什么？

# ①通信系统的基本概念



最简单的通信系统模型

三个基本单元：信源（**Source**）、信道（**Channel**）和信宿（**Destination**）。

# ①通信系统的基本概念

- 信源是提供消息的人、设备或事物。
  - 打电话时人发出的声音
  - 因特网上计算机系统中制作好的可供浏览的网页
  - 广播、电视发送的节目信号
  - 天体天电信号
- 人们最感兴趣的是**信源给出的消息**。
- **消息**是由符号、文字、数字、语声、图像等组成的序列。

# ①通信系统的基本概念

消息通常具有如下特征：

## (1) 消息可以产生、传递和获取

- 消息是客观的、可感知的，但它看不见、摸不着
- 传递要借助于载体（或媒质），用载体传送，从载体中取出
- 载体是消息的物理体现，如：
  - 在纸上写字，
  - 用磁带、光碟、存储器等记录语音、图像、数据
- 最值得研究的载体是电的、磁的、光的载体，这就是各类信号
- 信号是一种载体，是消息的物理体现，它使消息能够看得见，变得具体化
- 人们通过时域、频域及各种变换域来研究信号的波形、频谱等结构，构成比较完善的信号理论

# ①通信系统的基本概念

## (2) 消息是有内容的

- 消息的内容是客观存在的
- 对消息的接受者来说，消息有新、旧或已知、未知之分，有用、无用及有大用、小用之别
- 客观存在的消息，对不同的接收者来说，其性质、作用不同



促使人们去研究消息的内涵以及人的主观判断等主观因素，从而产生了信息理论及其各种分支。

# ①通信系统的基本概念

- 消息从信源通过信道到信宿，如何有效、可靠地传输，是通信系统要解决的根本问题。
  - 先说有效性（**Validity**）
    - 消息若在信源中先去粗取精，则必能提高通信的有效性
    - 对数字通信来说，即信源编码（**Source Coding**）要解决的主要问题
  - 再说可靠性（**Reliability**）
    - 信宿对接收到的“消息”若能够进行判断、评估，去伪存真的处理，则必能提高通信的可靠性
    - 对数字通信来说，信道编码（**Channel Coding**）的主要任务就是解决这个问题

# ①通信系统的基本概念

信息理论最初就是从解决通信系统的有效性和可靠性而发展起来的，从信源、信道、信宿的特性入手，逐步发展成应用极为广泛的新兴学科——信息科学



为此，本章首先给出信源、信道、信宿的比较确切的定义，然后再讨论信息的涵义。

# ①通信系统的基本概念

## ■ 信源：

- 消息的源，通常指向通信系统提供消息的人、设备或事物。
- 信源输出的是以符号形式出现的具体消息，它载荷信息。
- 信源输出的消息有多种形式：
  - 例如，从消息在时间上和幅度上的分布情况这一角度，可分为离散信源（**Discrete Source**）和连续信源（**Continuous Source**）。
  - 前者指信源发出的消息在时间和幅度上都是离散分布的，如字母、文字、数字等符号组成的符号序列或单个符号。
  - 后者指在时间和幅度上都是连续分布的，如话音、图像等。

# ①通信系统的基本概念

## ■ 信道：

- 传递消息的通道，广义上是指从信源到信宿间传递物理信号的媒质和设施。
- 信道的种类很多，如电信中常用的架空明线、同轴电缆、波导、光纤、传输电磁波的空间等都是信道。
- 也可以从信道的性质或其传送的信号情况来分类，例如：无干扰信道和有干扰信道、恒参信道和变参信道、离散信道（**Discrete Channel**）和连续信道（**Continuous Channel**）、单用户信道和多用户信道等。

# ①通信系统的基本概念

## ■ 信宿：

消息传递的对象，即接收消息的人、设备或事物。根据实际需要，信宿接收的消息，其形式可以与信源发出相同，也可以不同。当两者形式不同时，存在一个映射关系。

## ②信息科学的有关概念

### 信息的一般含义及定义

- 从通信的实质意义来讲，如果信宿收到的消息是已知的，则等于没有收到任何消息。
- 因此，人们更感兴趣的是消息中所包含的未知的成分，用概率论的术语来说，就是**具有不确定性的成分**。消息中这种未知的或不确定的成分，通常被称为消息中所包含的信息。
- 由此看来，通信中的信息、消息和信号是紧密相联的：
- 通信系统传输的是信号，信号是消息的载体，消息中的未知成分是信息。

## ②信息科学的有关概念

### 信息的一般含义及定义

- 教科书对消息和信号有明确的定义
- 但如何给信息下定义始终是个困扰大家的问题。
- 我国的《辞海》对信息作了如下解释：
  - 信息：通信系统传输和处理的对象，泛指消息和信号的具体内容和意义。通常须通过处理和分析来提取。**信息、物质和能量**被称为**系统的三大要素**。信息和物质的量值与信息的随机性有关，如在接收端无法预估消息或信号中所蕴含的内容或意义，即预估的可能性越小，信息量就越大。
- 此解释首先肯定了信息、消息和信号的密不可分性

## ②信息科学的有关概念

■ 即将出版的《科学技术大系》将这三者合在一起加以说明：  
信息·消息·信号：

□ **信息**是认识主体（人、生物、机器）所感受的或所表达的事物运动的状态和运动状态变化的方式。信息是人们在适应外部世界和控制外部世界的过程中，同外部世界进行交换的内容。信息的特征为：

- 接收者在收到信息之前，对它的内容是不知道的，所以信息是新知识、新内容；
- 信息是能使认识某一事物的未知性或不确定性减少的有用知识；
- 信息可以产生，也可以消失，同时信息可以被携带、贮存及处理；
- 信息是可以量度的，信息量有多少的差别。

## ②信息科学的有关概念

□ **消息**是指包含有信息的语言、文字和图像等。

- 消息中的不确定成分是信息。
- 在通信中，消息是指担负着传送信息任务的单个符号或符号序列。
- 这里的符号包括文本、数据、语言、图形和图像等。
- 消息是具体的，它载荷信息，但它不是物理性的。

□ **信号**是消息的物理体现，为了在信道上传输消息，就必须把消息加载到具有某种物理特征的信号上去。

- 信号是信息的载荷子或载体，它是物理性的。
- 在近代通信中信号往往是电信号和光信号。

## ②信息科学的有关概念

- 应该说，以上两种解释具有代表性和一定的权威性，都力图对信息给出正确及全面的定义。然而，信息科学还在飞速发展，人们对信息也不断赋予新的内涵，对信息的理解不但要考虑信息理论发展的历史，也要考虑它的将来，因此，  
➤ 正确及全面地理解它的涵义远比记住它的定义重要。
- 为此，先讨论信息科学的有关概念及信息理论研究的主要内容，再回过头来看看这些定义，也许会有更深的理解。

## ②信息科学的有关概念

### 信息与消息和信号的区别

- **消息：**是指包含有信息的语言、文字和图像等，可表达客观物质运动和主观思维活动的状态。
- 信息必须依附于一定的物质形式存在，这种运载信息的物质，称为**信息载体**。
- 消息中包含信息，是信息的载体。
- 同一信息可用不同的消息形式来载荷。
- 一则消息也可载荷不同的信息，可能包含丰富的信息，也可能包含很少的信息。
- 消息和信息既有区别又有联系。

## ②信息科学的有关概念

### - 信息与消息和信号的区别

- 信号: 把消息变换成适合信道传输的物理量, 这种物理量称为信号(如电信号、光信号、声音信号等)。
  - 信号携带着消息, 它是消息的运载工具。
  - 信号携带信息, 但不是信息本身。
  - 同一信息可用不同的信号来表示。
  - 同一信号也可表示不同的信息。

## ②信息科学的有关概念

### - 各种信息的概念

- 信息是信息论中最基本、最重要的概念，它是一个既抽象又复杂的概念。
- 信息的概念是在实践中产生，以前一直被看作是消息的同义词。
- 信息的概念十分广泛，不同的定义在上百种以上。
  - 信息是事物之间的差异；
  - 信息是事物联系的普遍形式；
  - 信息是物质的普遍属性。
  - .....
- 数学家认为：信息是使概率分布发生改变的东西。
- 哲学家认为：信息是物质的意识成分按完全特殊的方式融合起来的产物。

## ②信息科学的有关概念

### - 信息的直观认识

第一个重要概念：信道上传送的是随机变量的值。

注意：

- 这就是说，我们在收到消息之前，并不知道消息的内容。否则消息是没有必要发送的。
- 消息随机变量有一个概率分布。
- 消息随机变量的一个可能取值就称为一个事件。

## ②信息科学的有关概念

### - 信息的直观认识

第二个重要概念：事件发生的概率越小，此事件含有的信息量就越大。（不太可能发生的事件竟然发生了，令人震惊）

例 事件“中国足球队**5: 0**力克韩国足球队”含有的信息量大。（小概率事件发生了，事件信息量大）

例 事件“中国足球队**0: 1**负于韩国足球队”含有的信息量小。（大概率事件发生了，事件信息量小）

## ②信息科学的有关概念

### - 信息的直观认识

第三个重要概念：消息随机变量的随机性越大，此消息随机变量含有的信息量就越大。

例 消息随机变量 $X$ =“中国足球队与巴西足球队比赛的结果”，则消息随机变量 $X$ 含有的信息量小。

(随机性小，可预见性大，因此该消息随机变量含有的信息量小。)

例 消息随机变量 $X$ =“意大利足球队与德国足球队比赛的结果”，则消息随机变量 $X$ 含有的信息量大。

(随机性大，可预见性小，因此该消息随机变量含有的信息量大。)

## ②信息科学的有关概念

### - 信息的直观认识

第四个重要概念：两个消息随机变量的相互依赖性越大，它们的互信息量就越大（这里指的是绝对值大）。

例  $X=$ 武汉明日平均气温， $Y=$ 黄陂明日平均气温， $Z=$ 北京明日平均气温， $W=$ 纽约明日平均气温。则

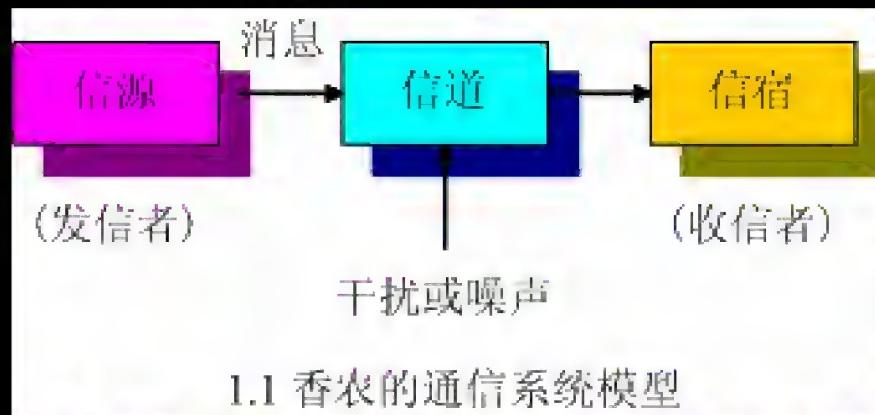
$X$ 与 $Y$ 互信息量大，

$X$ 与 $Z$ 互信息量小得多，

$X$ 与 $W$ 互信息量几乎为0。

### (3) 香农和维纳对信息的定义

- 香农：创立了信息论，但并没有给出信息的确切定义，他认为“信息是事物运动状态或存在方式的不确定性的描述”。



- 维纳：信息论的创始人之一，把人与外部环境交换信息的过程看作是一种广义的通信过程。后来才认识到“信息既不是物质又不是能量，信息就是信息”。即信息是独立于物质和能量之外存在于客观世界的第三要素。

# （4）信息的定义和性质

- ① 如何定义信息
- ② “本体论”层次
- ③ “认识论”层次
- ④ 香农信息论及其性质

# ① 如何定义信息

- 各种信息定义的差异：侧面不同、详略不同、抽象程度不同、概括层次高低不同。
- 根据不同的条件、区分不同的层次，可以给信息下不同的定义。

<http://www.survivor99.com/lcg/books/GIT/GY/>

## ② “本体论”层次

- 最高层次是最普遍的层次，也是无约束条件的层次。在这个层次上定义的信息是最广义的信息，使用范围也最广。
- 每引入一个条件，定义的层次就降低一点，使用的范围就变窄一点。
- “本体论”层次定义：信息是该事物运动的状态及其变化方式。

### ③ “认识论”层次

- 认识论层次：站在认识主体的立场上。即信息是认识主体（生物或机器）所感知的或所表述的相应事物运动状态及其变化方式。
- 虽然认识论比本体论的层次要低一些，所定义信息的使用范围也要窄一些，但是信息概念的内涵比本体论要丰富得多。
- 全信息：同时考虑外在形式/语法信息、内在含义/语义信息、效用价值/语用信息，称为全信息。
  - 语法信息：事物运动状态和状态改变的方式；
  - 语义信息：事物运动状态和方式的具体含义；
  - 语用信息：事物运动状态和方式及其含义对观察者的效用。

### ③“认识论”层次

- 认识论层次的信息是同时考虑语法信息、语义信息和语用信息的全信息。
- 本体论信息是第一位的，认识论信息是第二位的。前者从纯客观的立场来定义信息，后者从认识主体的立场来定义信息。
- 一般而言，在人类所及的有限时空中，本体论信息与认识论信息之间存在着交互作用和互相转化的关系，其过程与人类认识和改造世界的过程是一致的。

## ④

# 香农信息论及其性质

- 香农信息论层次：考虑事物运动状态及其变化方式的外在形式，实际上研究的是语法信息。
- 信息的定义：是对事物运动状态和变化方式的表征，它存在于任何事物之中，可以被认识主体（生物或机器）获取和利用。
- 信息的抽象性和重要性
- 信息的**5**条主要特征
- 信息的**11**条重要性质

## ■ 信息的抽象性和重要性

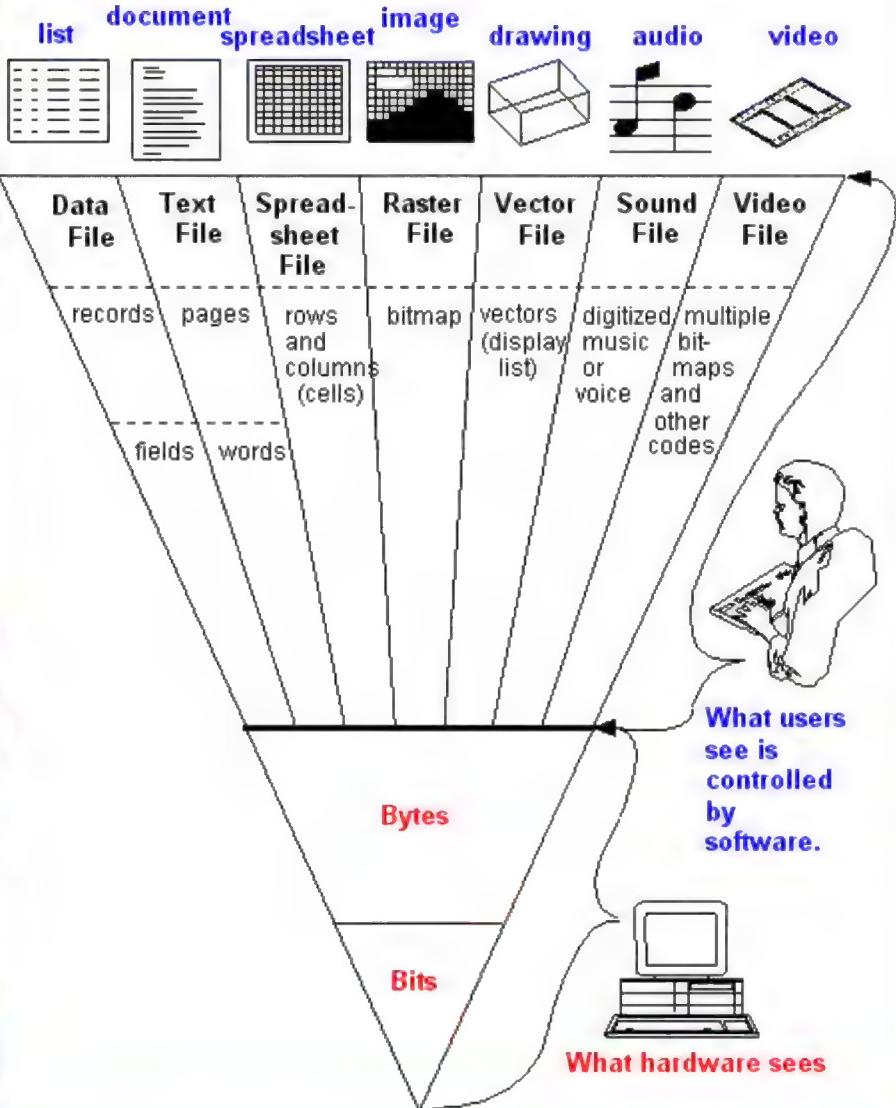
- **抽象性：**信息本身既看不见，又摸不着，没有气味、没有颜色、没有形状、没有大小、没有重量.....，它是非常抽象的东西。
- **重要性：**信息在信息化程度越来越高的社会中将起到越来越重要的作用，是比**物质**和**能量**更为宝贵的资源，全面**掌握**信息的概念，**正确**、**及时**、**有效地利用**信息，能够为人类创造更多的财富。

# Turn your data into information.



Information has “meaning”

Information  
Structure



## ■ 信息的5条主要特征

- 信息来源于物质，又不是物质本身；它从物质的运动中产生出来又可以脱离物质而相对独立地存在。
- 信息来源于精神世界，但又不局限于精神领域。
- 信息与能量息息相关，但又与能量有本质的区别。
- 信息具有知识的本性，但又比知识的内涵更广泛。
- 信息可以被认识主体获取和利用。

# ■ 信息的11条重要性质

- 存在的普遍性：信息存在于自然界、人类社会，其本质是运动和变化的。哪里有事物的运动变化，哪里就会产生信息。它处处存在，它既区别于物质和能量，又与物质和能量有相互依存的关系。
- 有序性：要使一个系统从无序变为有序，必须从外界获取信息。
- 相对性：同一个事物，不同的观察者获得的信息量可能不同。
- 可度量性：信息的多少用信息量表示。
- 可扩充性
- 可存储、传输与携带性
- 可压缩性：信息论研究的主要问题之一。
- 可替代性：可替代劳力、资本、物质材料、时间。
- 可扩散性：可在短时间内较大范围扩散开。
- 可共享性：没有信息的共享性就没有人类社会的发展和进步。
- 时效性：信息是有“寿命”的。

# 1.2 信息的分类

- 信息的定性认识和定量认识。
- 信息是一种非常复杂的研究对象，信息分类有许多不同的准则和方法。

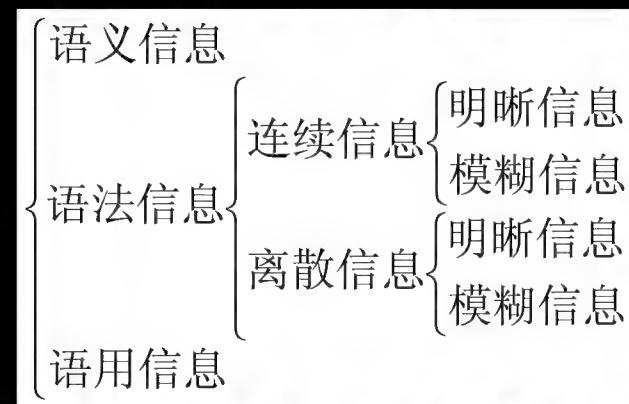
- (1) 信息的分类
- (2) 我们研究的内容
- (3) 香农信息论的缺陷

# (1) 信息的分类

- 按照信息性质：语法信息、语义信息、语用信息；
- 按照观察过程：实在信息、先验信息、实得信息；
- 按照信息地位：客观信息、主观信息；
- 按照信息作用：有用信息、无用信息、干扰信息；
- 信息逻辑意义：真实信息、虚假信息、不定信息；
- 信息传递方向：前馈信息、反馈信息；
- 信息生成领域：宇宙信息、自然信息、社会信息、思维信息；
- 信息应用部门：工业信息、农业信息、军事信息、政治信息、科技信息、文化信息、经济信息、市场信息、管理信息；
- 按照信息来源：语声信息、图像信息、文字信息、数据信息、计算机信息；
- 信息载体性质：语声信息、图像信息、文字信息、电磁信息、光学信息、生物信息；
- 携带信息的信号性质：连续信息、离散信息、半连续信息。

## (2) 我们研究的内容

- 最重要的分类方法是按照信息性质的分类
- 信息性质分类的三个基本类型



- 按照事物的运动方式，信息可进一步分为概率信息、偶发信息、确定信息、模糊信息；还可以分为有限状态和无限状态。香农信息论主要研究语法信息中的**概率信息**。

### (3) 香农信息论的缺陷

- ① 香农定义信息的出发点是假定事物状态可以用一个经典集合论为基础的概率模型来描述。实际存在的某些事物运动状态要寻找一个合适的概率模型往往是非常困难的。某些情况下，这一模型不一定存在。
- ② 香农定义信息的度量没有考虑收信者的主观性和主观意义，也撇开了信息的具体含义、具体用途、重要程度和引起的后果等元素。这就与实际情况不完全一致。

香农信息的定义和度量是科学的，能反映信息的某些本质；但却是有缺陷的、有局限的。这样，它的适用范围会受到严重影响。

# 1.3 信息论的起源、发展及研究内容

- (1) 信息传输手段的五次重大变革
- (2) 信息论的形成和发展
- (3) 信息论的研究对象、目的和内容
- (4) 香农信息论的科学体系
- (5) 本课程的内容
- (6) 信息科学的发展

# (1) 信息传输手段的五次重大变革

- ① 语言产生：人们用语言准确地传递感情和意图，使语言成为传递信息的重要工具。
- ② 文字产生：人类开始用书信的方式交换信息，使信息传递的准确性大为提高。
- ③ 发明印刷术：使信息能大量存储和大量流通，并显著扩大了信息的传递范围。
- ④ 发明电报电话：开始了人类电信时代，通信理论和技术迅速发展。这一时期还诞生了无线电广播和电视。更深入的问题：如何定量研究通信系统中的信息，怎样更有效、更可靠传递信息？
- ⑤ 计算机与通信结合：促进了网络通信的发展，宽带综合业务数字网的出现，给人们提供了除电话服务以外的多种服务，使人类社会逐渐进入了信息化时代。

## (2) 信息论的形成和发展

- 1948年：信息论学科诞生：1948年，香农发表论文“*A mathematical theory of communication*”。同时维纳提出了最佳滤波理论，成为信息论的一个重要分支。
- 五十年代：信息论在学术界引起了巨大反响。  
1959年：香农为各种信源编码的研究奠定基础：发表论文“*Coding theorems for a discrete source with a fidelity criterion*”，数据压缩理论与技术成为信息论的重要分支。
- 六十年代：信道编码技术有较大进展。信道编码成为信息论重要分支。  
1961年：香农的重要论文“双路通信信道”开拓了多用户信息理论的研究。
- 七十年代：有关信息论的研究，从点对点间的单用户通信推广到多用户系统的研究。密码学成为信息论重要分支。信息安全的重要性。
- 目前的发展

## (2) 信息论的形成和发展

### ■ 信息论的创始人

- 信息科学的创立是以**信息理论**（简称**信息论**）的建立为标志的。
- 美国科学家香农（C. E. Shannon）在1948年和**1949**年先后发表的两篇文章“**The mathematical theory of communication**”和“**Communication in the presence of noise**”，奠定了信息论的基础。
- 香农在这两篇论文中，讨论了信息的度量、信息特征、信息传输速率、信道容量、干扰对信息传输的影响等问题，**全部理论建立在信息是可以度量的基础上**，但他没有给出信息的定义，只是提出问题，分析问题，得出结论。

## (2) 信息论的形成和发展

- 美国另一科学家维纳 (N. Wiener) 出版了“**Extrapolation, Interpolation and Smoothing of Stationary Time Series**”和“**Control Theory**”这两本名著。
  - 维纳是控制论领域的专家，重点讨论微弱信号的检测理论，并形成了信息论的一个分支。
  - 他对信息作了如下定义：信息是人们在适应外部世界和控制外部世界的过程中，同外部世界进行交换的内容的名称。

## (2) 信息论的形成和发展

- 信息是不是物质？有没有能量的属性？与物质、能量的关系是什么？
- **Shannon**和**Wiener**没有给出上述问题的解答，随着信息理论向通信以外的学科特别是向社会科学领域的渗透，人们为了弄清这些问题进行了很多研究，但直到现在也还没有定论。

## (2) 信息论的形成和发展

一般认为信息论的创始人是香农和维纳，但由于香农的贡献更大，所以更多的人认为香农更合适。



- 维纳，美国数学家，控制论的创始人。1894年11月26日生于密苏里州的哥伦比亚，1964年3月18日卒于斯德哥尔摩。
- 维纳在其50年的科学生涯中，先后涉足哲学、数学、物理学和工程学，最后转向生物学，并且在各个领域中都取得了丰硕的成果，称得上是恩格斯颂扬过的、本世纪多才多艺和学识渊博的科学巨人。他一生发表论文240多篇，著作14本，自传两本《昔日神童》和《我是一个数学家》。
- 维纳的主要成果有八个方面：建立维纳测度、引进巴拿赫—维纳空间、位势理论、发展调和分析、发现维纳—霍普夫方法、提出维纳滤波理论、开创维纳信息论、创立控制论。



# 目前的发展

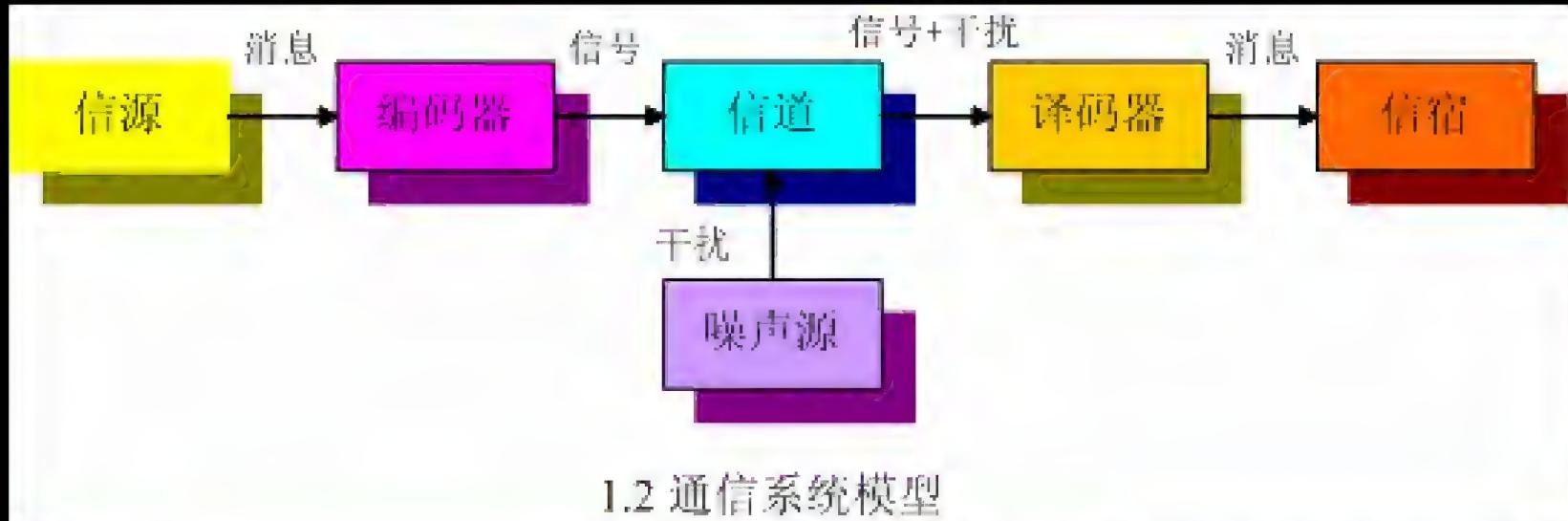
- 香农信息论：
  - 信息概念的深化；
  - 网络信息理论和多重相关信源编码理论的发展和应用；
  - 通信网的一般信息理论研究；
  - 信息率失真理论的发展及其在数据压缩和图像处理中的应用；
  - 信息论在大规模集成电路中的应用；
  - 磁记录信道的研究等。
- 纠错码理论：在工程方面应用及最优编码方法研究。
- 维纳信息论：对量子检测和估计理论、非参数检测和估计理论以及非线性检测与估计理论的研究。

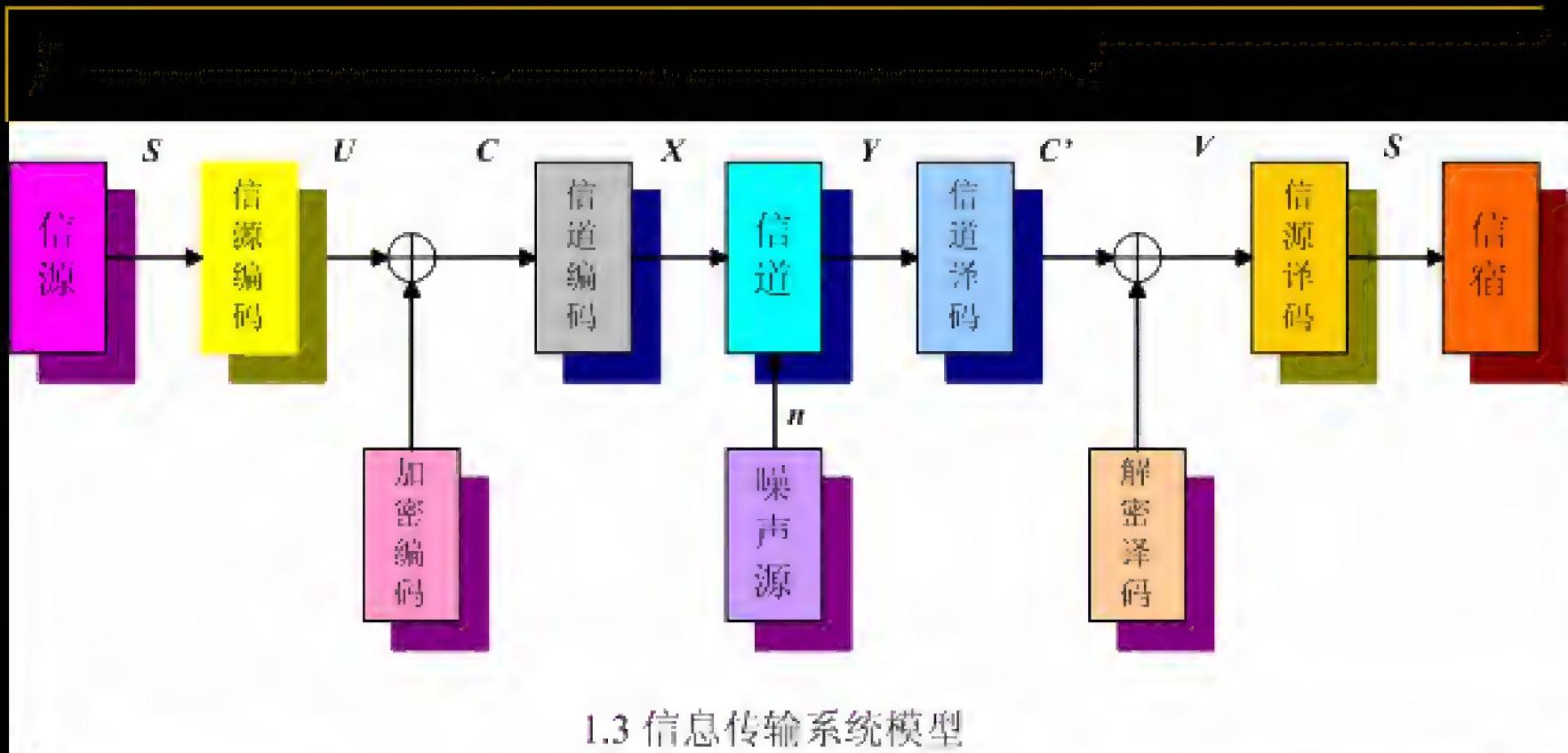
### (3) 信息论研究的对象、目的和内容

- ① 信息论研究的对象和目的
- ② 信息论研究的内容
- ③ 香农和维纳理论的区别

# ① 信息论研究的对象和目的

- 研究对象：信息论是一门应用概率论、随机过程、数理统计和近世代数的方法，来研究广义的信息传输、提取和处理系统中一般规律的工程学科。





### 1.3 信息传输系统模型

- 研究目的：它的主要目的是提高信息系统的可靠性、有效性和安全性以便达到系统最优化。

## ② 信息论的研究内容

### ■ 信息论要回答的一些最基本的理论问题：

- 什么是信息？如何度量？
- 在给定的信道中，信息传输有没有极限？
- 信息能否被压缩和恢复？极限条件是什么？
- 从实际环境中（如干扰、噪声）抽取信息，极限条件是什么？
- 在允许一定失真（**Distortion**）的条件下，信息能否被更大程度地压缩？极限条件是什么？
- 设计什么样的系统才能达到上述极限？
- 现实中，接近极限的设备是否存在？

## ② 信息论的研究内容

- 信息论的成功就在于首先对这些问题给出了明确的回答，这就是信息量、信道容量、熵（**Entropy**）、信息率失真函数、**Shannon**的三个定理和**Shannon**公式等。
- 事实上，回答上述问题只是信息论要解决的问题中的一部分，是建立在**Shannon**研究成果的基础上，被称为**Shannon**信息论或狭义信息论。

## ② 信息论的研究内容

- **信息论基础：**也称狭义信息论/经典信息论/香农信息论。主要研究信息测度、信道容量、信息率失真函数，与这三个概念相对应的香农三定理，信源编码，信道编码。
- **一般信息论：**主要研究信息传输和处理问题。除香农基本理论之外，还包括噪声理论、信号滤波和预测、统计检测和估计理论、调制理论。后一部分内容以维纳为代表。
- **广义信息论：**是一门综合性的新型学科—信息科学。至今没有严格的规定。凡是能够用广义通信系统模型描述的过程或系统，都能用信息基本理论来研究。不仅包括一般信息论的内容，还包括如医学、生物学、心理学、遗传学、神经生理学、语言学、语义学、社会学、经济管理中有关信息的问题。反过来，所有研究信息的识别、控制、提取、变换、传输、处理、存储、显示、价值、作用和信息量的大小的一般规律以及实现这些原理的技术手段的工程学科，都属于广义信息论的范畴。

## ② 信息论的研究内容

### 狭义信息论

- 狹义信息论主要总结了Shannon的研究成果，因此又称为Shannon信息论。
  - 在信息可以度量的基础上，研究如何有效、可靠地传递信息。
- 有效、可靠地传递信息必然贯穿于通信系统从信源到信宿的各个部分，
- 狹义信息论研究的是收、发端联合优化的问题，而重点在各种编码。
- 它是通信中客观存在的问题的理论提升。

## ② 信息论的研究内容

### 一般信息论

- 研究从广义的通信引出的基础理论问题：
  - **Shannon**信息论；
  - **Wiener**的微弱信号检测理论。
    - 微弱信号检测又称最佳接收
    - 是为了确保信息传输的可靠性，研究如何从噪声和干扰中接收信道传输的信号的理论。
    - 主要研究两个方面的问题：从噪声中去判决有用信号是否出现和从噪声中去测量有用信号的参数。
    - 该理论应用近代数理统计的方法来研究最佳接收的问题，系统和定量地综合出存在噪声和干扰时的最佳接收机结构。

## ②

# 信息论的研究内容

## 一般信息论

- 这里的“接收机”只是对受到噪声与干扰“淹没”及破坏的信号进行数学处理，主要目的是将其设计为“最佳”。
- 所谓“最佳”总是同某些特定的假设条件及某个准则相联系的，因此这里的最佳接收机是在给定条件下，能够最好满足所给准则的接收机。在通信系统中常用的有最小差错准则、最大似然准则最小均方误差准则等。
- 信号检测具体又分为确知信号的检测和具有随机信号参量的信号的检测等，其重点在接收端，研究如何从干扰中提取信息。

## ② 信息论的研究内容

### 一般信息论

- 除此之外，一般信息论的研究还包括：噪声理论、信号滤波与预测、统计检测与估计理论、调制理论、信号处理与信号设计理论等。
- 可见它总结了Shannon 和Wiener以及其他学者的研究成果，是广义通信中客观存在的问题的理论提升。

## ② 信息论的研究内容

### 广义信息论

- 无论是狭义信息论还是一般信息论，讨论的都是客观问题。然而从前面给出的例子可知，当讨论信息的作用、价值等问题时，必然涉及到主观因素。
- 广义信息论研究包括所有与信息有关的领域，如：心理学，遗传学，神经生理学，语言学，社会学等。
- 因此，有人对信息论的研究内容进行了重新界定，提出从应用性、实效性、意义性或者从语法、语义、语用方面来研究信息，分别与事件出现的概率、含义及作用有关，其中意义性、语义、语用主要研究信息的意义和对信息的理解，即信息所涉及的主观因素。

## ② 信息论的研究内容

### 广义信息论

- 广义信息论从人们对信息特征的理解出发，从客观和主观两个方面全面地研究信息的度量、获取、传输、说是很多问题的理论复杂，于人类知识论目前还过于人本制于正在发展的阶段。

### ③ 香农和维纳理论的区别

- 香农理论：香农研究的对象是从信源到信宿之间的全过程，是收、发端联合最优化问题，其重点是放在编码。他指出，只要在传输前后对消息进行适当的编码和译码，就能保证在干扰的存在下，最佳地传送和准确或近似地再现消息。为此发展了信息测度理论、信道容量理论和编码理论。
- 维纳理论：维纳研究的重点是在接收端。研究一个信号如果在传输过程中被某些因素（如噪声、非线性失真等）所干扰时，在接收端怎样把它恢复、再现，从干扰中提取出来。在此基础上，创立了最佳线性过滤理论（维纳滤波器）、统计检测与估计理论、噪声理论。

# (4) 香农信息论的科学体系

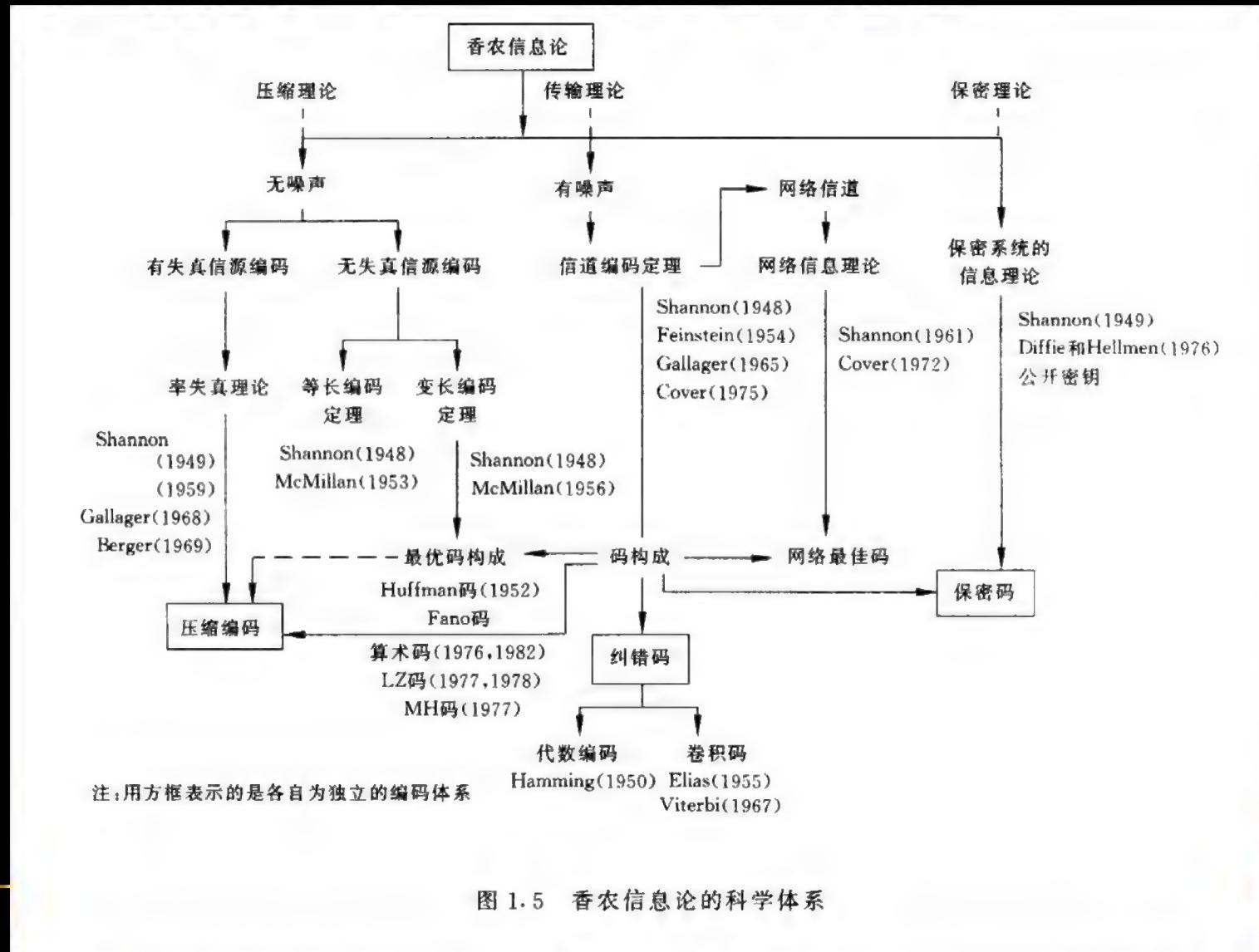


图 1.5 香农信息论的科学体系

# (5) 本课程的内容

- 信息论基础—香农信息论。
  - 主要包括
    - 信息测度—信源熵
    - 信道容量
    - 信息率失真函数
    - 香农三定理：
      - 无失真信源编码定理
      - 信道编码定理
      - 限失真信源编码定理
    - 信源编码
    - 信道编码
    - 保密信息理论

# (6) 信息科学的发展

- 信息科学已成为世界各国最优先发展的科学之一。
- 在我国，信息科学更是受到前所未有的重视，这对信息科学工作者来说都是极大的鼓舞。
- 本小节将对信息科学发展背景、通信工程与基础信息结构、信息理论与信息系统的关系、信息学科、信息学科介绍，简单的学关系，使读者明确认息理论的主动性和自觉性。

# (6) 信息科学的发展

## ■ 信息科学的定义及迅速发展的背景

- 信息科学是研究信息的获取、存储、传输、加工处理和利用等内容的技术科学。
- 它与数学、物理、材料、生命、心理等基础学科以及众多技术学科交叉形成新领域，是发展和更新最为迅速的技术科学研究领域之一。
- 信息科学与其它科学技术交叉将派生出大量的新概念、新构思、新技术和边缘学科。
- 其核心是现代通信加信息技术。

# （6）信息科学的发展

## ■ 国家信息基础结构

- 亦称“信息高速公路”，是一个具有大容量、高速度的电子数据传输系统，最早由美国联邦政府引导工业界提出并逐步建立，简称NII（National Information Infrastructure）。
- 它以现代信息技术为基础，以光导纤维网络为骨干，连接全美的学校、研究机构、企业、图书馆、实验室乃至每个家庭，使人们共享丰富的信息资源。

## (6) 信息科学的发展

- 一般认为，NII由通信网、计算机、各种终端设备、信息和人等5个基本部分组成，它的主要特点是：
  - 能以极快的速度高度准确地传送数字化的音像；
  - 以新的方式存储、传输和处理大量的数据；
  - 是一个开放的系统，通过现有的电话线、有线电视接口或其它通用接口，进入每个人的生活以共享信息资源。

# (6) 信息科学的发展

## ■ 信息产业

- 信息产业在世界范围内正在由先导产业逐步变为主导产业。
- 从微观上看，表现为单位产品的价格构成中，能源和材料的消耗减少而信息技术和信息服务的比重上升；
- 从宏观上看，表现为国民生产总值（GDP）中信息产业所占的比重增加。
- 一个国家信息产业的发展水平将是衡量该国社会经济总体发展和现代化程度的重要标志之一。

# （6）信息科学的发展

## ■ 信息与通信工程一级学科

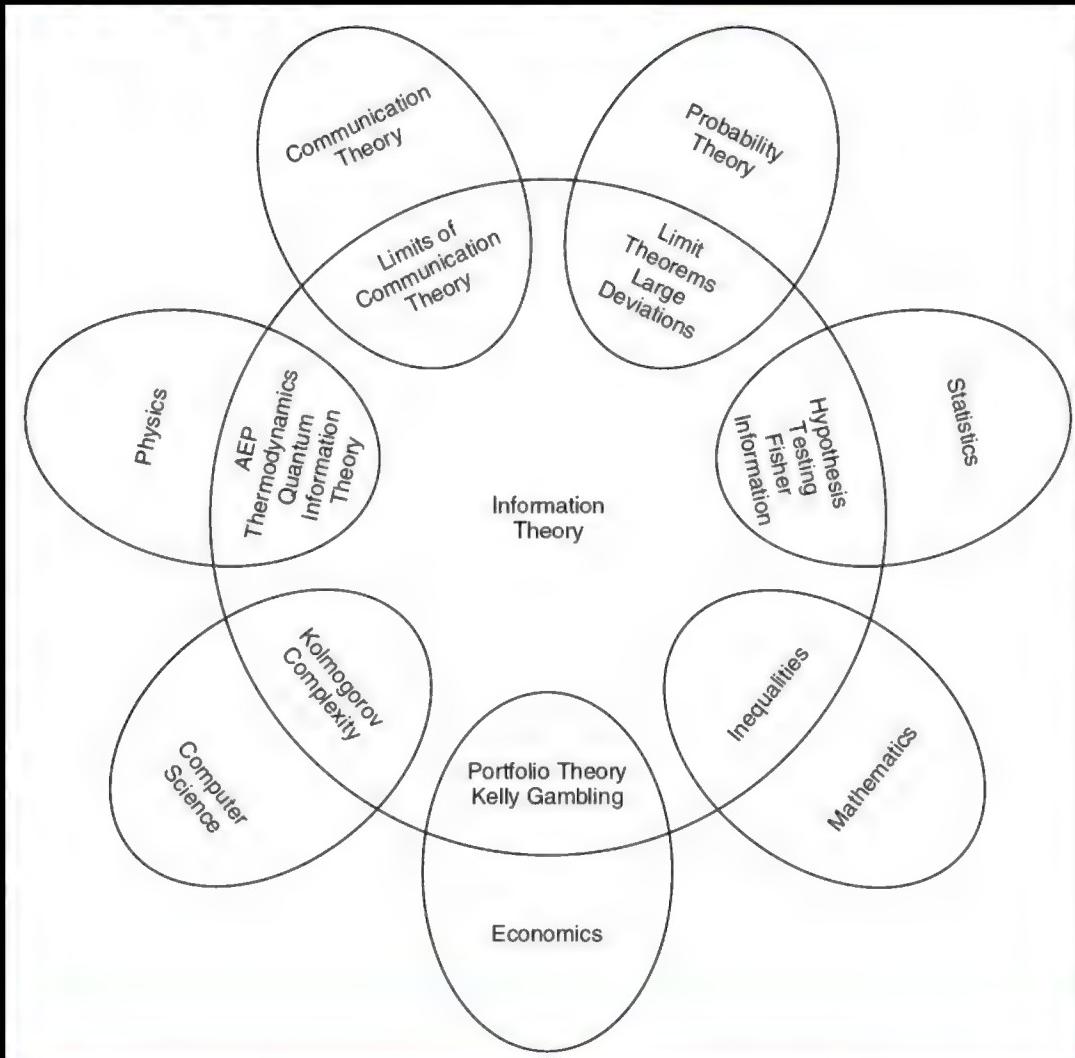
- 九十年代中期，国家在进行科学的研究的学科规划和调整时，将工程学科原来**100**多个一级学科调整为**31**个，但却将原来的“电子与信息系统”调整为“信息与通信工程”和“电子科学与技术”两个一级学科。
- 信息与通信工程下设“通信与信息系统”和“信号与信息处理”两个二级学科。

# (6) 信息科学的发展

- 信息理论是信息科学的基础
- 强调用数学语言来描述信息科学中的共性问题及解决方案
- 这些共性问题分别集中在狭义信息论、一般信息论和广义信息论中

# (6) 信息科学的发展

与其它学科的联系



# 概率论回顾

随机变量 $X, Y$ 分别取值于集合 $\{x_1, x_2, \dots, x_i, \dots, x_n\}$ 和 $\{y_1, y_2, \dots, y_i, \dots, y_m\}$ :

$$(1) \quad 0 \leq p(x_i), p(y_j), p(x_i / y_j), p(y_j / x_i), p(x_i y_j) \leq 1$$

$$(2) \quad \sum_{i=1}^n p(x_i) = 1, \sum_{j=1}^m p(y_j) = 1, \sum_{i=1}^n p(x_i / y_j) = 1,$$

$$\sum_{j=1}^m p(y_j / x_i) = 1, \sum_{j=1}^m \sum_{i=1}^n p(x_i y_j) = 1$$

# 概率论回顾

$$(3) \quad \sum_{i=1}^n p(x_i y_j) = p(y_j), \sum_{j=1}^m p(x_i y_j) = p(x_i)$$

$$(4) \quad p(x_i y_j) = p(x_i)p(y_j / x_i) = p(y_j)p(x_i / y_j)$$

$$(5) \quad \text{当 } X \text{ 与 } Y \text{ 相互独立时, } p(y_j / x_i) = p(y_j), \\ p(x_i / y_j) = p(x_i), p(x_i y_j) = p(x_i)p(y_j)$$

$$(6) \quad p(x_i / y_j) = \frac{p(x_i y_j)}{\sum_{i=1}^n p(x_i y_j)}, p(y_j / x_i) = \frac{p(x_i y_j)}{\sum_{j=1}^m p(x_i y_j)}$$

# Suggested Reading

- Rafael Capurro, Birger Hjørland, "The Concept of Information", Annual Review of Information Science and Technology Ed. B. Cronin, Vol. 37 (2003) Chapter 8, pp. 343-411.

[http://www.db.dk/bh/publ\\_uk.htm](http://www.db.dk/bh/publ_uk.htm)

- Robert M. Losee, "A Discipline Independent Definition of Information", J. of the American Society for Information Science 48 (3) 1997, 254-269.

<http://ils.unc.edu/~losee/b5/book5.html>

- "Doubts about the concept of information"- An elementary derivation of the Shannon entropy formula ,

<http://okmij.org/ftp/Computation/limits-of-information.html>

- 鲁晨光 著,<<广义信息论>>, 中国科技大学出版社, 1993·合肥

<http://www.survivor99.com/lcg/books/GIT/GY/>

# 思考题

一位朋友很不赞成“通信的目的是传送信息”及“消息中未知的成分就算是信息”这些说法。他举例说：我多遍地欣赏梅兰芳大师的同一段表演，百看不厌，大师正在唱的正在表演的使我愉快，将要唱的和表演的我都知道，照你们的说法电视里没给我任何信息，怎么能让让我接受呢？请从信息论的角度对此做出解释。（主要从狭义信息论与广义信息论研究的内容去理解和解释）

# Thank you!

Quiz Time!

